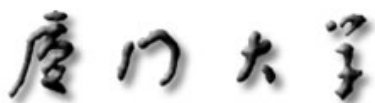


学校编码: 10384

分类号_____密级

学号: X2012230594

UDC



工 程 硕 士 学 位 论 文

基于数据仓库的银行客户信息管理系统的设计与实现

Design and Implementation of Customer Information
Management System for Banking Industry Based on Data
Warehouse

林 乐

指 导 教 师: 夏侯建兵副教授

专 业 名 称: 软 件 工 程

论文提交日期: 2014 年 2 月

论文答辩日期: 2014 年 4 月

学位授予日期: 2014 年 月

指 导 教 师:

答辩委员会主席:

2014 年 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1.经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ☒ ） 2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

建立以客户为中心的服务体系，对于银行维持客户资源，争取新客户，挖掘潜力客户，增强市场竞争力有着重要意义。在目前我国商业银行中，尽管各银行在数据管理系统的建立方面投入规模庞大，但大多数还处于比较低的水平，与世界发达国家的银行系统有着很大的差距。银行数据挖掘是从大型数据库中发现并提取所需信息的过程，目的是帮助银行分析人员寻找各种数据之间的关联和规律，从而提供有效的决策支持。

本文提出了一种基于数据仓库、数据挖掘的银行客户关系管理系统解决方案，该数据仓库从多个信息源中获取原始数据，经数据转换清洗、整理加工后，存储在数据仓库的内部数据库中，通过数据仓库访问工具，向数据仓库的用户提供统一、协调和集成的信息环境，支持企业全局的决策过程和对企业经营管理的深入综合分析，以此来作为对客户提供服务和产品营销的辅助手段。

本文立足银行业的实际需求，运用日益成熟的软件开发技术成果和网络技术，基本实现了符合银行业客户管理需求和其他相关工作要求解决方案。对于银行业的信息化建设进一步规范化流程提供了一定的参考依据，具有相对较高的使用价值，基本能够满足各家银行的客户管理和其他当前经营管理需要。

关键词：数据仓库；数据挖掘；银行客户信息管理

Abstract

In order to strive for new customers excavate the potential customers and enhance market competitiveness, establishing a service system which centered on customers will have great significance. In the current, although our commercial banks put a large scale of investment on data management, our skills are still in low level and have a large gap with developed countries. Bank data digging is a procedure of finding and extract information, which aims to help bank analysts to find out the rules and similarities between the various data so as to provide effective decision support.

This dissertation puts forward a kind of banking customer relationship management system solution based on data warehouse and data mining, which to obtain the original data from multiple sources through cleaning, finishing, storing in the internal database. Provide the user of the data warehouse unity, coordination and integration of information environment, support enterprises in global decision-making process and further comprehensive analysis of the enterprise management, in order to for customers to provide services and products marketing means of auxiliary.

This dissertation based on the actual needs of the banking industry, the use of increasingly sophisticated software development and technological achievements and network technology, achieved banking customer relationship management needs in line banking and other related work required solutions. Provide some reference for the banking industry information construction more standardized processes, has a relatively high value, and can basically meet the bank's current banking customer relationship management and other management needs.

Key words: The Data Warehouse; Data Wining; Bank Customer Information Management

目 录

| | |
|-----------------------------|-----------|
| 第一章 绪论 | 1 |
| 1.1 研究背景与研究意义 | 1 |
| 1.1.1 研究背景 | 1 |
| 1.2.1 研究意义 | 1 |
| 1.2 国内外研究现状 | 2 |
| 1.2.2 国外研究现状 | 2 |
| 1.2.3 国内研究现状 | 3 |
| 1.3 论文的研究内容 | 4 |
| 1.4 论文的组织结构 | 5 |
| 第二章 系统开发相关技术介绍 | 6 |
| 2.1 数据仓库 | 6 |
| 2.1.1 数据仓库概念 | 6 |
| 2.1.2 OLAP | 7 |
| 2.1.3 数据仓库技术 | 8 |
| 2.2 数据挖掘 | 9 |
| 2.2.1 数据挖掘概念 | 9 |
| 2.2.3 数据挖掘环境 | 11 |
| 2.3 本章小结 | 12 |
| 第三章 系统需求分析 | 13 |
| 3.1 业务需求 | 13 |
| 3.2 用户需求 | 14 |
| 3.2.1 客户价值内涵 | 14 |
| 3.2.2 系统用例分析 | 18 |
| 3.3 功能需求分析 | 20 |
| 3.3.1 关键营运指标功能 | 22 |
| 3.3.2 客户营销分析功能 | 22 |
| 3.3.3 产品营销分析功能 | 23 |
| 3.3.4 数据挖掘报告功能 | 23 |
| 3.4 非功能需求分析 | 23 |
| 3.5 本章小结 | 25 |

| | |
|----------------------------|-----------|
| 第四章 系统概要设计 | 26 |
| 4.1 系统设计目标与原则 | 26 |
| 4.1.1 系统设计目标 | 26 |
| 4.1.2 系统原则 | 26 |
| 4.2 系统功能模块设计 | 27 |
| 4.3 数据仓库系统构架 | 28 |
| 4.3.1 数据来源描述 | 28 |
| 4.3.2 数据仓库部分 | 30 |
| 4.3.3 数据聚集部分 | 31 |
| 4.3.4 数据展示部分 | 35 |
| 4.4 系统架构设计 | 36 |
| 4.5 系统数据库设计 | 38 |
| 4.5.1 数据表组织 | 38 |
| 4.5.2 系统应用 | 39 |
| 4.6 系统安全设计 | 43 |
| 4.7 本章小结 | 44 |
| 第五章 系统详细设计与实现 | 45 |
| 5.1 系统设计界面实现 | 45 |
| 5.2 系统各模块实现 | 47 |
| 5.2.1 关键运营指标模块 | 47 |
| 5.2.2 客户营销分析模块 | 53 |
| 5.2.3 产品营销分析模块 | 58 |
| 5.2.4 数据挖掘报告模块 | 63 |
| 5.3 本章小结 | 65 |
| 第六章 总结与展望 | 66 |
| 6.1 总结 | 66 |
| 6.2 展望 | 67 |
| 参考文献 | 68 |
| 致 谢..... | 70 |

Contents

| | |
|-------------------------------------------------------------------|-----------|
| Chapter 1 Introduction..... | 1 |
| 1.1 Research Background and Significance | 1 |
| 1.1.1 Research Background | 1 |
| 1.1.2 Research Significance | 1 |
| 1.2 Research Actuality | 2 |
| 1.2.1 The Developing Status Abroad | 2 |
| 1.2.2 The Developing Statusat Home | 3 |
| 1.3 Main Contents..... | 4 |
| 1.4 Organization of the Dissertation | 5 |
| Chapter 2 Introduction to System Related Technologies..... | 6 |
| 2.1 Data Warehouse..... | 6 |
| 2.1.1 Data Warehouse | 6 |
| 2.1.2 OLAP | 7 |
| 2.1.3 Data Warehouse Technology..... | 8 |
| 2.2 Data Mining | 9 |
| 2.2.1 Concept of Data Mining | 9 |
| 2.2.3 Data Mining Environment | 11 |
| 2.3 Summary | 12 |
| Chapter 3 System Analysis..... | 13 |
| 3.1 The Business Requirements..... | 13 |
| 3.2 The User Needs | 14 |
| 3.2.1 Customer Value Connotation | 14 |
| 3.2.2 System Use Case Aanalysis | 18 |
| 3.3 FunctionalRequirements..... | 20 |
| 3.3.1 Key Operation Index Function | 22 |
| 3.3.2 Customer Marketing Analysis Fnction | 22 |
| 3.3.3 Product Marketing Analysis Capabilities | 23 |
| 3.3.4 Data Mining Report Fnction | 23 |
| 3.4 Non-functional Requirements | 23 |
| 3.5 Summary | 25 |

| | |
|----------------------------------------------------------|-----------|
| Chapter 4 Summary System Design | 26 |
| 4.1 System Design Objectives and Principles | 26 |
| 4.1.1 System Design Goals | 26 |
| 4.1.2 System Policy | 26 |
| 4.2 System function module design..... | 27 |
| 4.3 Data Warehouse System Architecture | 28 |
| 4.3.1 Data Source Description | 28 |
| 4.3.2Data Warehouse | 30 |
| 4.3.3Data Aggregation | 31 |
| 4.3.4Data Presentation | 35 |
| 4.4 System Architectures | 36 |
| 4.5 System Database Design | 38 |
| 4.4.1 Data Table Organization | 38 |
| 4.4.2 System Application | 39 |
| 4.6 System Database Design | 43 |
| 4.7 Summary..... | 44 |
| Chapter 5 Detailed Design And Implementation..... | 45 |
| 5.1 Interface to Realize the System Design | 45 |
| 5.2 The System Module..... | 47 |
| 5.2.1 Key Operation Index Function | 47 |
| 5.2.2 Key Operation Index Function | 53 |
| 5.2.3 Product Marketing Analysis Capabilities | 58 |
| 5.2.4 Data Mining Report Fnction | 63 |
| 5.3 Summary..... | 65 |
| Chapter 6 Conclusions and Outlook | 66 |
| 6.1 Conclusions..... | 66 |
| 6.2 Outlook | 67 |
| References..... | 68 |
| Acknowledgments | 70 |

第一章 绪论

1.1 研究背景与研究意义

1.1.1 研究背景

日趋激烈的市场竞争使得银行对于数据处理提出了更高的要求,他们已经不满足于信息系统能够迅速地处理业务,而是需要从浩如烟海的大量数据中探索业务活动的规律,提炼出经营管理所必须的关键信息,使银行能对其自身业务的运作以及整个市场相关行业的态势进行分析,从而进行正确决策,并最终通过市场竞争转变为利润。虽然我国银行业已经积累了大量的历史数据,事务处理型技术在业务领域的应用也已经趋于成熟^[1]。但相对而言,分析处理型技术的研究和应用还处于起步阶段。对于现有业务数据的分析处理,大多数银行还只是仅限于做出一些统计报表,至于如何进一步发掘这些报表中反映的信息,则需要依靠分析人员付出大量的劳动进行“人工获取”。但有时对于某些深藏在其中的有价值的规律,通过“人工方式”可能永远都无法发现,这远没有达到支持决策的要求。因此,银行迫切需要一种集成方案来解决业务数据“过剩”而决策所需信息贫乏的局面。

目前正被各个领域广泛应用的数据挖掘技术就能较好的解决银行所面临的问题。所谓数据挖掘,简单地说就是对一些看似无规则的海量数据进行整理,采用某些算法或模型处理,发现数据间内在联系或推导出有价值的结论。因此,数据挖掘技术的发展为银行的科学决策提供了很好的技术支持。根据这种现状,作者对数据挖掘技术在银行中的一些具体应用进行了较为全面的归纳和深入的并通过数据挖掘技术对银行各种信用卡数据进行了分析处理,得出了一些要的规律和结论,并据此提出了建设性的管理建议。以此希望能为国内银理决策和数据挖掘应用提供一些有益的启示。

1.2.1 研究意义

本文提出了一种基于数据仓库、数据挖掘的银行客户关系管理系统解决方案,较好地构建一个标准的数据库仓库,通过有效充分的数据挖掘,将银行客户

资料作不同角度的分析，从中对客户进行定位分类，明确客户的消费倾向与消费模式，预测客户的风险性和利润性，以此来作为对客户提供服务和产品营销的辅助手段。运用一系列先进、完整的银行客户数据模型，对不同层次、不同类型客户进行细分，集客户分析、营销管理、考核管理等诸多分析和应用功能于一体的自动智能化的新型电子化运营支撑系统，具有相对较高的使用价值，基本能够满足各家银行的客户关系管理和其他当前经营管理需要。

1.2 国内外研究现状

1.2.2 国外研究现状

国外的日本三和银行是引入客户关系管理的重要代表，并成立了专门的客户关系管理部门，向其他网点或子公司提供全面、系统的金融服务，客户关系管理成为其保持国际竞争力的有力武器。除了在管理制度重视客户关系管理之外，还要建立相应的客户关系管理信息系统，挖掘客户需求。

数据仓库是金融机构实现客户关系管理(Customer Relationship Management)的核心技术，也是金融业竞争优势的来源。主要的应用业务部门为信用卡部、信贷部、市场部和零售业务部等，应用领域是以客户为中心的分销渠道管理、客户利润分析、客户关系优化和风险控制与管理。这些外资银行面对全球化这么复杂的经营环境，对外可以做好客户关系管理，对内能有效控制风险，除了经营团队的专业能力外，最主要的原因是采用数据仓库技术建立以客户为中心的管理信息系统，领导可以科学化手段分析海量的历史数据进行决策。目前全球排名前二十大的银行中，有超过三分之二的银行已经采用数据仓库技术作为以客户为中心的管理信息系统平台，平均的获利率超过 12%。

1. 美洲银行^[2]

通过实施数据仓库，实现了更好的客户保留和客户获取、增强交叉销售的效果、优化市场营销活动效率、大大提供了目标销售的成功率。详细情况如下：

(1) 数据仓库的应用为 Bank of America 带来价值\$40,000,000 的新业务、风险避免和利润。

(2) 以事件为基础的营销活动的成功率达 40% 以上。

2.加拿大皇家银行

通过实施数据仓库，大幅度提高了客户细分的精确度和准确性，在目标客户营销活动中，RSP 储蓄有 11%的提高，实现了以价值为基础的客户分析，增加了 20%的高利润贡献客户。从而使得它们能够在过去的三年中，每年市场营销带来的收入都有两位数字的增长。详细情况如下：

- (1) 与原来的客户细分相比，有接近 75%的客户发生了迁移；
- (2) 在营销活动的周期上提高 61%；
- (3) 提高高价值客户的数目&平均收益&贡献水平(13%)；
- (4) 提高直接市场营销的反馈率>(40%)等。

1.2.3国内研究现状

国内金融行业也充分认识到这一方面的差距，从 2000 年开始，国内几大商业行和地方、民营的金融机构已纷纷开始了数据仓库的研究，都在着手调研、准备或者尝试实施基于数据仓库技术的各种解决方案，一些银行已进入了实施阶段。中国工商银行进行了以个人客户关系管理(PCRM)和业绩价值管理 P(VMS)为主的应用试点；中国银行则全面规划了信用卡系统，其中很重要的一个子系统就基于数据仓库技术的销售和客户服务系统；中国农业银行正在广东分行进行经分析系统的建设；民生银行也全面启动了客户信息管理(CIM)和企业级数据仓库建设。但是，国内各行数据仓库这几年的建设结果却差强人意，鲜有成功的案例。往往是“雷声大，雨点小”，也就是说，项目的初始设定目标与最终实现结果相甚远。那么问题究竟出在哪里呢？概括起来主要表现在三个方面：

1.片面追求先进性，缺乏对银行自身的经营体制、管理水平、数据的可获得等诸多现实情况的深入分析，项目初始目标定得太大，片面追求大而全，其结局必然是欲速则不达；

2.系统缺乏可扩展性，往往是初始项目能够取得一定效益，但是后续项目很难在原有基础上进行扩展，系统的开发应用变得越来越复杂，有时迫不得已另起炉灶，造成资源的巨大浪费；

3.重业务系统的开发，轻数据源分析整理、数据标准化、数据模型设计、体系结构设计等基础工作，造成数据仓库项目基础不牢，“好看不好用”。

1.3 论文的研究内容

1.本文在对国内外银行数据仓库的应用现状进行了深入论述和分析的基础上,结合银行实际,给出了一套数据仓库系统的总体设计方案和实施的具体方法。

2.采取了数据仓库总线模型构架,使得数据仓库开发团队可以按照总线规范异步地独立地建设各个数据库集市,该构架可以缩短建设周期和费用,具有良好的一致性和可拓展性。

3.数据抽取、转换和加载过程非常复杂,本文提出了数据以文件形式传输并建立数据集中区的方案,可以最大程度的减少对业务的干扰,适合银行的现有 IT 环境。

4.通过对客户的交易行为、持有产品等进行分析,挖掘出目标客户,为管理者制定销售计划和方案提供支持,支持一线人员对目标客户的主动销售,使产品销售更加具有针对性;

5.运用系统的方法,按照客户的特点和需求把客户市场划分成若干具有共同特征的客户子市场,为针对客户的特点采取独特的产品或市场营销战略,提供差异化的产品和服务提供数据分析支持。

6.挖掘产品交叉销售的目标客户群体,以及进行交叉销售的产品组合(套餐),从而实现对存量客户直接进行产品的交叉销售;对增量客户在客户办理业务时可按照系统分析结果提供产品套餐,从而提高产品销售能力;

7.利用技术工具,通过对客户进行分析,建立中高端客户的提升分析模型,选择重点产品(理财、基金)建立客户购买意向分析模型,充分挖掘出客户的潜在价值;

8.为各级个人业务的管理人员提供与客户、产品相关的业务管理报表和灵活的查询统计功能;

9.为总行高管层和个人业务部门负责人提供个人业务经营管理关键指标查看功能,及时提供全行个人经营关键指标进展情况;为高层管理人员及时了解业务现状,及时调整经营策略提供参考。

1.4 论文的组织结构

本文共分六章，各章安排如下：

第一章绪论，介绍项目的研究背景及研究意义，以及银行的国内外发展现状，最后介绍论文的研究内容及组织结构。

第二章系统相关技术介绍，重点介绍数据仓库、数据挖掘等技术。

第三章系统需求分析，详细介绍了银行业客户管理系统的分析过程，包括系统的可行性分析、功能性需求、非功能性需求以及功能性需求的一些具体用例。

第四章系统概要设计，详细介绍了银行业客户管理系统的系统设计原则、设计架构以及该系统的运行环境，在此同时，详细介绍了客户管理系统的数据库设计。

第五章系统详细设计与实现，详细介绍了银行客户管理系统的系统界面的实现、系统各个模块的实现。

第六章总结和展望，对论文的主要工作内容进行了总结，并提出未来应扩展的方向进行了阐述。

第二章 系统开发相关技术介绍

2.1 数据仓库

面对当今日趋激烈的竞争市场,各级管理人员迫切需要从企业信息系统中获取一致性的、有效的决策支持信息,及时准确地把握市场变化的脉搏,做出正确有效的判断和抉择。也就是说,数据处理的重点应该从在线业务处理扩展到在线分析处理。人们一直试图利用传统数据库技术来支持在线分析处理,但效果一直不能令人满意。这是因为事务处理和分析处理有着根本不同的性质,数据仓库作为企业应用集成中数据集成层面的重要技术在金融行业得到广泛应用。

2.1.1 数据仓库概念

数据仓库是“面向主题的、集成的、稳定的、随时间变化的数据集合,用于支持管理决策过程”。在这个定义中,W.H.Inmon^[3]明确给出了数据仓库的四个重要特点,即“面向主题”“集成”“稳定”“随时间变化”,对数据源中的数据(包括运作数据库、外部市场数据、办公数据、档案数据等)进行收集,并对其进行标准化、过滤、净化等处理,盖上时间戳后,进入数据仓库(Data Warehouse),再经由各种工具(可以是OLAP工具、报表工具、DSS工具、数据挖掘工具等)对数据仓库中的数据进行知识发现,并将其应用于实践,为客户进行科学的决策提供理论支持。

数据仓库的数据组织方式主要包括有三种,分别是虚拟存储方式、基于关系表的存储方式和多维数据库存储方式等其中,多维数据库存储方式是直接面向数据挖掘分析操作所需的数据组织形式,它对DW中的海量数据从客户感兴趣的角度进行层次化处理、抽象概括,并设置维索引及相应的元数据管理文件,以对应于数据仓库中的数据与虚拟存储方式、基于关系表的存储方式中组织关系都比较复杂相比,更适用于组织、存储数据仓库中的海量数据。

2.1.2 OLAP

OLAP (On line Analytical Processing)的概念是关系数据库之父 E. F. Codd 于 1993 年提出的。他认为联机事务处理(OUT)已不能满足用户基于分析的综合查询和决策支持分析的要求。基于分析的综合查询和决策分析往往涉及到企业范围内某个主题数据的查询分析，用关系数据库的 SQL 查询来实现将涉及到大量表的连接计算，不仅非常复杂，而且其效率根本满足不了需求^[4]。因此，他提出了多维数据库和多维分析的概念，即 OLAP。

OLAP 是一类软件技术，如图 2.1 所示，它可以使各层次的用户通过采用多种可能的观察形式，对数据仓库中的数据数据进行快速、稳定、交互式的存取操作^[5]。其最显著特征是能够提供数据的多维概念视图。OLAP 的其它特性分别是快速响应用户的分析需求、完善的分析功能、安全的共享性、获取信息的及时性。

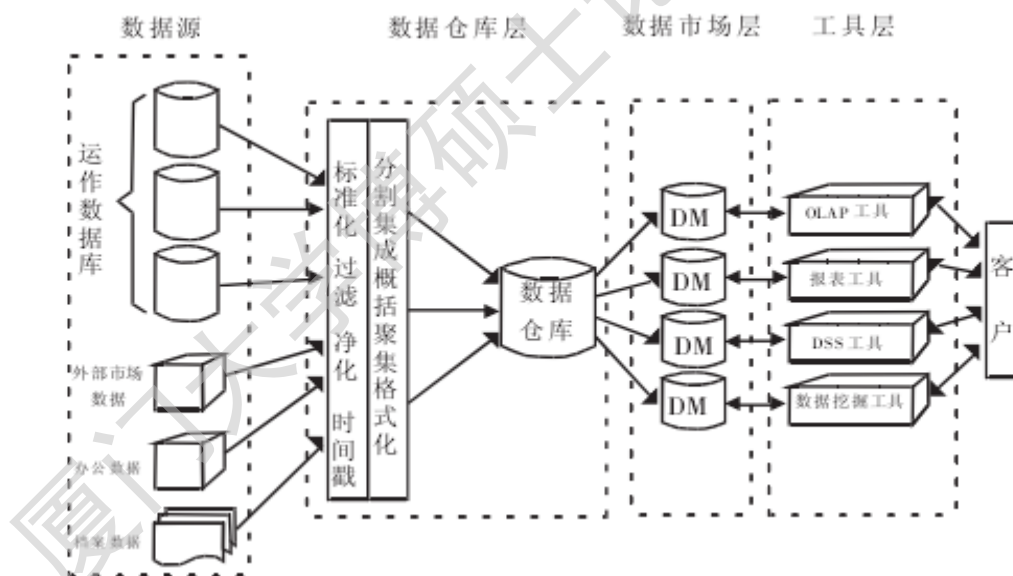


图 2.1: OLAP 技术

2. OLAP 多维分析操作

多维分析是指对数据仓库中的数据进行切片、切块、钻取和旋转等分析操作^[5]，用户能从多角度对数据进行深入剖析，进而了解其中的规律。

(1) 切片。切片是指通过在某个或某些维上选取某一特定的属性成员，而在其他维上则选取一定区间的属性成员，对所获得的数据进行分析。以某银行部分业务数据所形成的数据立方为例进行说明：从该立方中选取 2014 年 1 月这一

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”. Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库